The American Journal of Interdisciplinary Innovations and Research
ISSN 2642-7478 | Open Access

Type Original Research
PAGE NO. 106-122
DOI 10.37547/tajiir/Volume07Issue07-10

Evolving Architectures and Long-Horizon Planning in Multi-Agent Conversational Ai: A Decade in Review

Rohan Mandar Salvi
University of Maryland, Baltimore County, Arbutus, Maryland, United States

Pronob Kumar Barman
University of Maryland, Baltimore County, Arbutus, Maryland, United States

OPEN ACCESS

SUBMITED 18 June 2025
ACCEPTED 25 June 2025
PUBLISHED 27 July 2025
VOLUME Vol.07 Issue 07 2025

CITATION
Rohan Mandar Salvi, & Pronob Kumar Barman. (2025). EVOLVING ARCHITECTURES AND LONG-HORIZON PLANNING IN MULTI-AGENT CONVERSATIONAL AI: A DECADE IN REVIEW. The American Journal of Interdisciplinary Innovations and Research, 7(07), 106–122. https://doi.org/10.37547/tajiir/Volume07Issue07-10

Abstract- This systematic review surveys advances in conversational AI from 2015 to 2025, focusing on the emergence of modular multi-agent architectures, hierarchical reinforcement learning, and self- evolving agents. A quantitative synthesis of 63 studies indicates that memory-augmented, long- horizon planners improve task success rates by approximately 30% over flat policies, while meta- learning and lifelong learning approaches halve sample complexity in data-scarce domains. Despite these gains, current systems remain brittle under distribution shifts, lack principled safety guarantees, and provide few benchmarks for diagnosing co-adaptive failure modes in mission-critical applications.

Keywords: Multi-Agent Systems, Conversational AI, Adaptive Dialogue, Hierarchical Planning, Reinforcement Learning, Meta-Learning, Emergent Communication, Self-Evolving AI

1. Introduction

1.1 Background

Over the past decade, AI has progressively transformed conversational systems from simple rule-based interaction engines into sophisticated agents capable of maintaining coherent and human-like dialogue. As the complexity of real-world problems increases, the need for collaborative agent systems, in which each agent possesses specialized knowledge and capabilities, is increasing. This has led to the emergence of *Adaptive Multi-Agent Conversational AI* (AMACAI), a paradigm in which artificial agents interact with both users and each other, learning, evolving, and making autonomous decisions through multi-turn conversations [1].

Unlike traditional systems, AMACAI agents are equipped with adaptive reasoning, long-horizon planning, and self- evolution capabilities. These agents are capable of collaborative behaviors, dynamically sharing information, and adjusting strategies in response to changing conversational contexts [2]. Applications range from virtual assistants and collaborative robotics to smart tutoring systems and distributed-support platforms.

Despite significant strides in NLP and reinforcement learning, the integration of architectural design, real-time planning, and self-evolution mechanisms in multiagent systems remains underexplored. Understanding the interplay between these components is critical for developing intelligent, responsive, and scalable conversational agents.

1.2 Aim and Objectives

The primary objectives of this review are as follows.

- To assess architectural frameworks employed in adaptive multi-agent conversational systems.
- To evaluate the effectiveness of long-horizon planning strategies.
- To examine mechanisms that support agent self-evolution.
- To identify open research challenges and future directions.

1.3 Research Questions

This study seeks to answer the following research question:

1. What architectural designs are most prevalent in AMACAI, and how do they influence agent coordination and dialogue generation?
2. How is long-horizon planning implemented in multi-agent dialogue systems, and which techniques enhance coherence over extended interactions?
3. What mechanisms support the self-evolution of agents, and how do they affect learning efficiency, adaptability, and task success?
4. What limitations currently hinder the development of scalable and safe AMACAI systems, and what strategies can address these limitations?

1.4 Research Rationale

As the demand for complex and context-aware dialogue agents increases, single-agent systems reveal critical limitations in terms of flexibility, scalability, and situational awareness. Adaptive multi-agent systems offer a promising alternative by distributing intelligence across coordinated agents capable of joint decision-making [3]. However, the existing literature often treats architectural design, planning, and self-evolution in isolation. This review aims to synthesize these dimensions into a unified framework that can inform future research and practical implementations.

2 Literature Review

2.1 Introduction

Conversational AI has developed considerably over the last decade owing to advances in deep learning, natural language processing (NLP), and reinforcement learning. Conversational agents have never been as fluent and coherent as they are now, with the emergence of large language models (LLMs) such as GPT, PaLM, and Claude [4]. However, these developments have focused mostly on single-agent systems and have limited capabilities for dynamic collaboration, distributed cognition, and real-time adaptation.

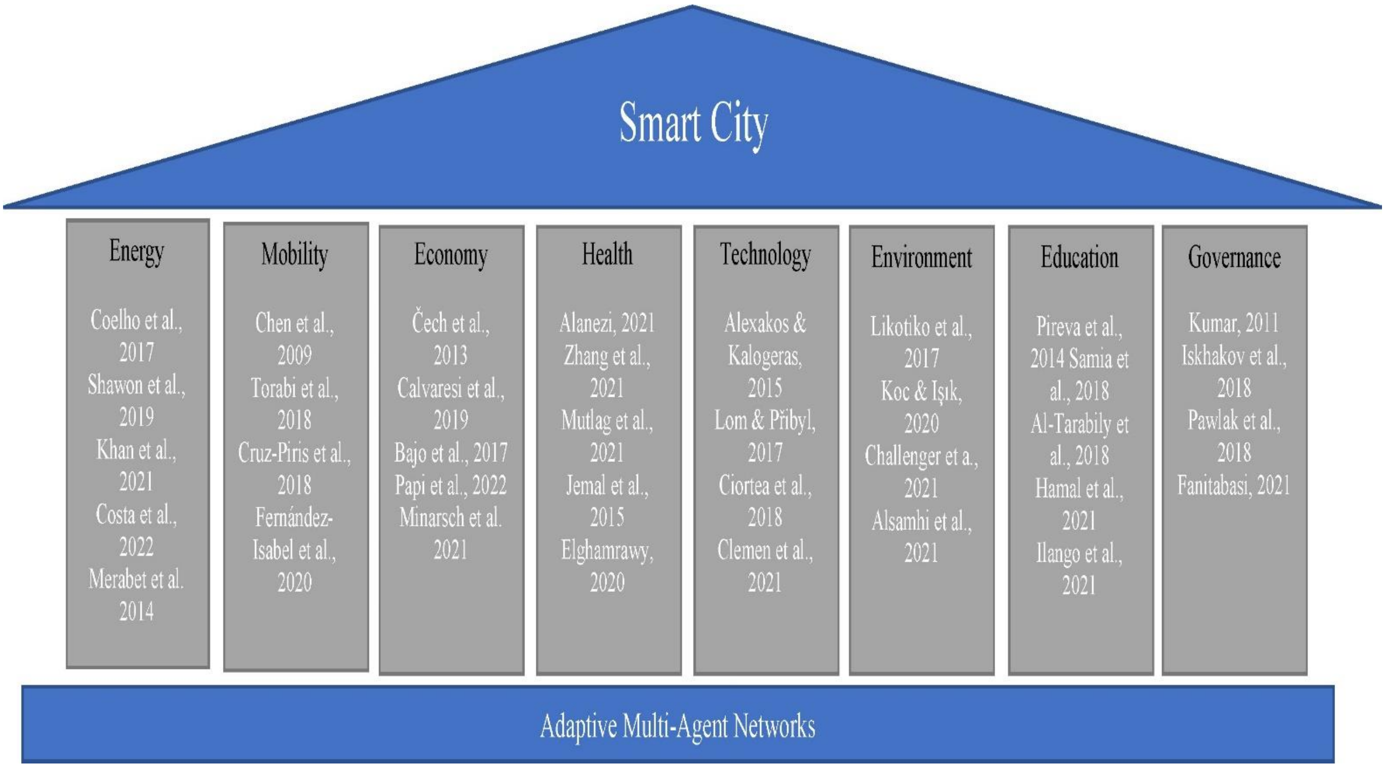

Figure 1: Adaptive Multi-Agent Networks [5]

Figure 1, which represents an adaptive multi-agent network, supports this discussion by illustrating the concept of multiple interacting agents.

The figure shows interconnected nodes, likely representing individual AI agents that form a complex network structure. This visual representation aligns with the text's description of AMACAI as a paradigm shift involving multiple specialized agents that interact and adapt over time. The network structure in the figure emphasizes the distributed and collaborative nature of AMACAI systems, in contrast to the single-agent approach mentioned earlier.

Adaptive Multi-Agent Conversational AI (AMACAI) is a conversational paradigm shift that uses multiple interacting agents, each specialized in various tasks, to have intelligent and context-aware conversations that adapt over time [6]. The literature review covers the background knowledge, architectural support, planning, and evolutionary mechanisms underlying this field of study. It also reviews the available research gaps and the wider scope of the study that scholars can explore in the future.

2.2 Literature Concept

2.2.1 History of Conversational Artificial Intelligence Systems

Early conversational systems were mostly rule-based and used scripted dialogues. Early systems, such as ELIZA and ALICE, are examples that show a basic understanding of language but cannot be flexibly applied to different situations or contexts. With the advancement of artificial intelligence, sequence-to-sequence neural models have emerged, playing a significant role in enabling systems to produce fluent and adaptive responses [1]. Attention mechanisms subsequently advanced the situational relevance of dialogue acts by enabling systems to emphasize significant sections of prior conversations. Even with such advancements, single-agent models have problems remaining coherent in prolonged interactions, and they do not necessarily adapt well to dynamic goals and changing user requirements.

2.2.2 Multi-Agent Systems in AI

Multi-agent systems (MAS) have their roots in the distributed artificial intelligence research area and were created to allow the computation of complex tasks with the cooperation of several entities. MAPle in dialogue systems, MAS can be used to delegate various conversational functions to dedicated agents. These roles include intent recognition, knowledge retrieval, emotional involvement, and dialogue planning skills. Each agent is typically characterized by a set of abilities or areas of knowledge that allow them to deal more effectively with multidimensional and complex dialogue situations. This separation of labor not only makes the system scalable but also makes the interactions more diverse, in addition to being deeper, because collaborative decisions can be made [1].

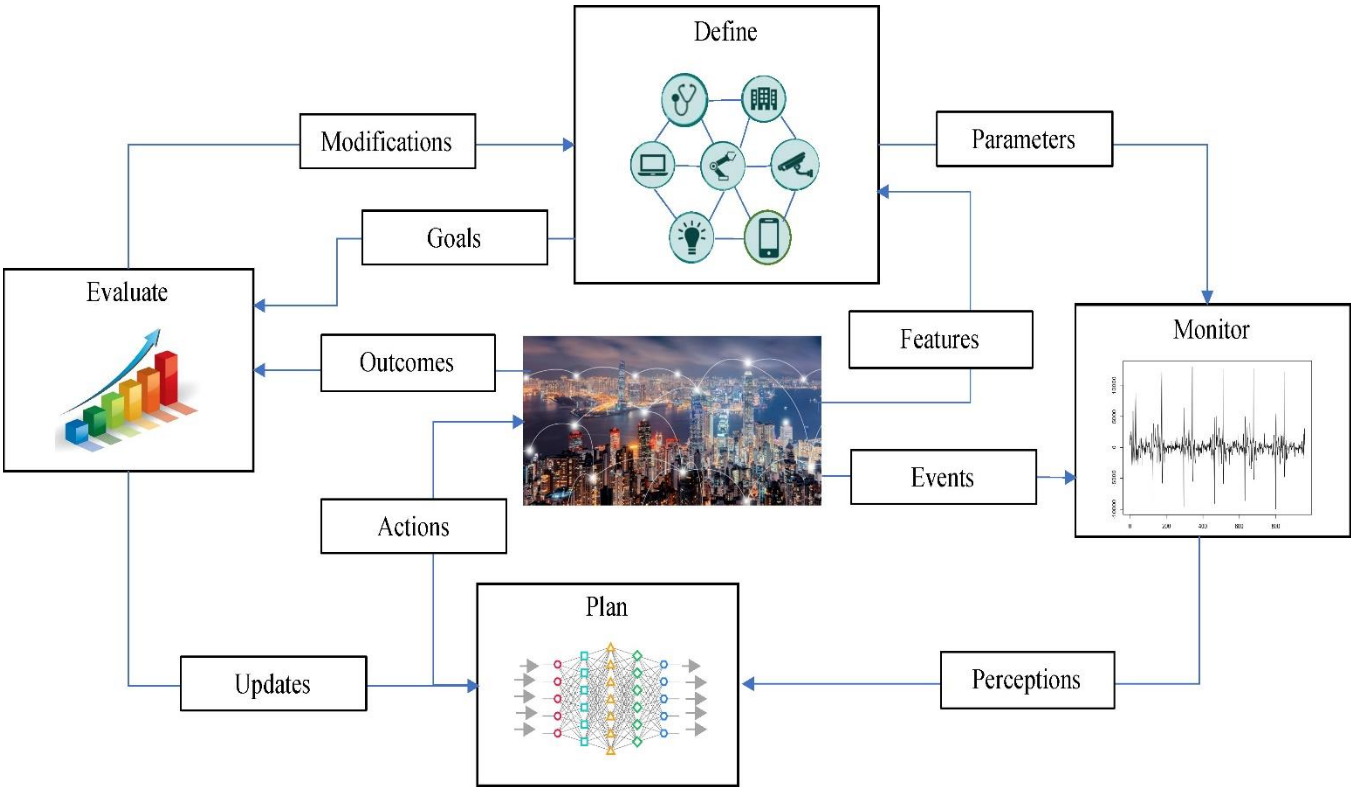

Figure 2: Overview of self-adaptive MAS [5]

Figure 2 provides an overview of self-adaptive multi-agent systems (MAS), as presented by Nezamoddini and Gho- lami [5]. The figure illustrates the key components and processes involved in the self-adaptive MAS framework.

The diagram shows a cyclical process with four main stages.

1. **Monitoring**: This stage involves collecting data from the environment and the system itself.
2. **Analysis**: The collected data is analyzed to detect any changes or issues that require adaptation.
3. **Planning**: Based on the analysis, the system plans the necessary adaptations or responses.
4. **Execution**: The planned adaptations are implemented, affecting both the system and its environment.

These four stages form a continuous feedback loop, allowing the MAS to constantly adapt to the changing conditions and requirements of the user.

2.2.3 Dialogue Management Using Adaptive Systems

One aspect of enhancing the conversation AI user experience is its adaptability. Adaptive systems change their responses depending on many factors, including user behavior, the history of a conversation, goals, and user preferences. The modification of dialogue strategies in real time involves reinforcement learning, probabilistic modeling, and user profiling techniques used in these systems. Adaptivity is distributed in a multi-agent configuration, and agents can learn not only through interactions with users but also through each other [2]. Coordinate adaptation is made possible with shared memory and inter-agent feedback loop mechanisms that help provide more coherent and context-dependent interactions. This feedback and constant learning enable the system to improve over time and become more helpful in various conversations.

2.2.4 Dialogue Long-Horizon Planning

Long-horizon planning describes the quality of a system that maintains dialogue context and intent during long conversations that are not interrupted. Rather than answering questions one at a time, long-horizon planning systems maintain an awareness of high-level objectives and can direct the flow of conversation in response to them [3]. Hierarchical planning and decision-making processes are techniques that aid such systems in breaking down long-term goals into smaller and achievable subtasks. In a multi-agent setting, these planning tasks can be shared among various agents, each dealing with particular parts or turns of the conversation. This modularity in planning enables the system to arrange more high-level and contextually relevant responses in the long run, so that users' objectives are addressed in an integrated manner.

Figure 3: Understanding Agentic AI: Attributes, Architecture, and the Ecosystem [7]

Figure 3 illustrates the key components and attributes of agentic AI systems, which are crucial for implementing long-horizon planning in dialogue systems. The figure showcases the interconnected nature of various AI technologies, including natural language processing, machine learning, and knowledge representation, all of which contribute to the development of sophisticated conversational agents capable of maintaining context and pursuing long-term objectives in dialogues [7]. This ecosystem approach highlights how different AI components work together to enable more coherent and goal-oriented conversations, aligning with the principles of long-horizon planning discussed in the context of dialogue systems.

2.2.5 Self-Develop

Self-evolution in conversational AI refers to allowing agents to continuously learn and enhance their performance independently of human intervention. Self-evolving systems contrast with static models that must be manually updated periodically to reflect new interactions and environmental feedback. Meta-learning and continual learning are two learning methods that enable such agents to generalise knowledge over tasks and learn new domains fast. In addition, emergent communication, in which agents create their own language or signaling systems during interactions, is superior to collaborative problem-solving and coordination [8]. In the long run, these self-improving abilities will create more personalization, strength to emerging challenges, and a more human-like development of conversational capabilities.

2.3 Theoretical Framework

An interdisciplinary combination of theories in cognitive science, artificial intelligence, communication studies, and control systems engineering forms the basis for the development of Adaptive Multi-Agent Conversational AI (AMACAI) systems. This theoretical background provides a conceptual representation of the plan of intelligent agents that can cooperate, adapt, and evolve in real-time conversational situations [6]. The theoretical constructs on which the AMACAI research is based are as follows:

2.3.1 Distributed Cognition

Distributed cognition is a cognitive science theory that states that cognitive activity is not bound to one person but is distributed among people, tools, and the environment. This notion, as far as AMACAI systems are concerned, is represented by the assignment of certain tasks to various agents in the system. All agents participate in global cognitive responses by performing specific information retrieval, planning, and sentiment analysis tasks. Together, the agents constitute a distributed network that permits complex reasoning and decision-making capabilities that are difficult to achieve by individual agents [9].

2.3.2 Multi-Agent Reinforcement Learning (MARL)

Multi-agent reinforcement learning is a variant of conventional reinforcement learning that is modified for application in multiagent environments. Such agents acquire policies through their interactions in a common environment and modify their behavior through trial-and-error or collaboration [10]. AMACAI systems often use system frameworks, such as centralized training and decentralized execution, which enable optimal group behavior and agent autonomy. This arrangement allows agents to easily draft strategies, react to dynamic responses, and optimize dialogue results in multilateral interactions.

2.3.3 Theory of Mind (ToM)

ToM is the ability to reason and anticipate the mental processes of others, including their beliefs, intentions, and desires. This theoretical view is critical in AMACAI systems to enable agents to model and act in response to other agents' or users' behaviors during conversations. Agents can produce more appropriate and contextually relevant dialogue by simulating the goals and possible reactions of other people. This increases the flow of interactions, especially in collaborative or multi-turn situations, where the important factor is to predict the behavior of the partner to remain in line with the goal [11].

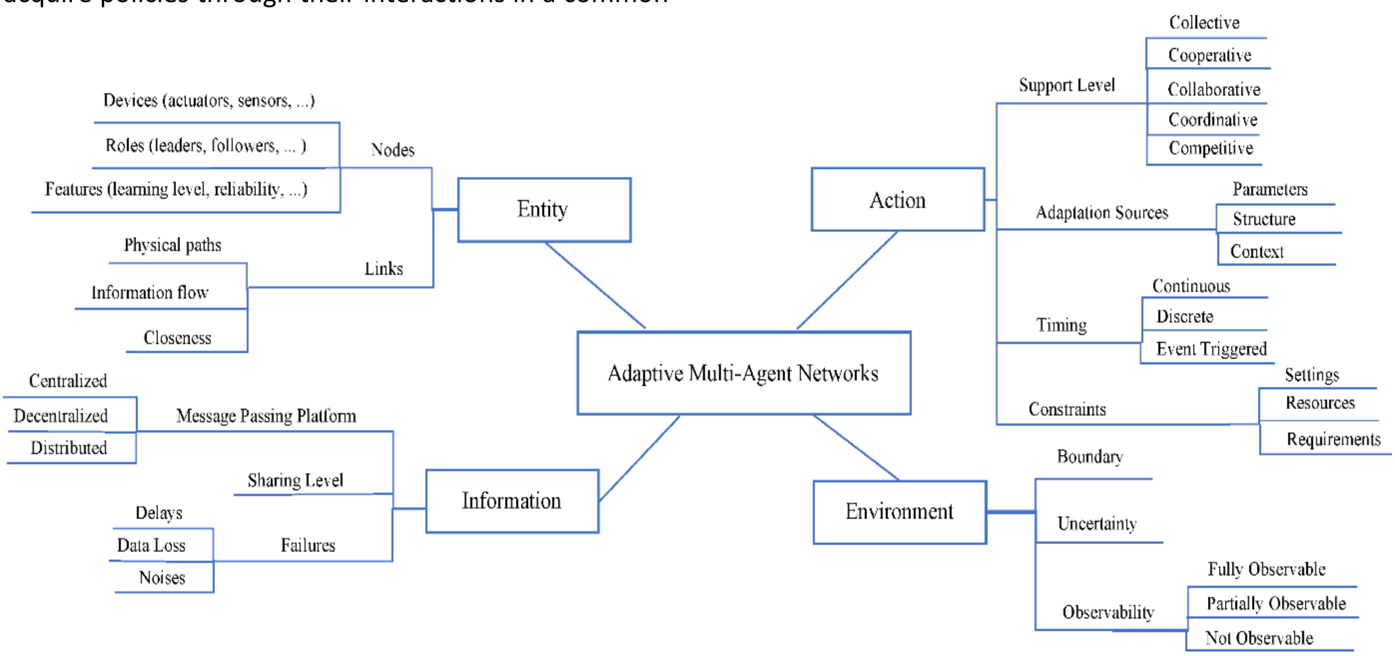

Figure 4: Contributing factors in defining adaptive MASs [5]

The importance of *Theory of Mind* is particularly evident in collaborative or multi-turn scenarios, where predicting the behavior of conversation partners is essential for maintaining alignment with the overall goal of the conversation. This ability contributes significantly to the adaptability of multi-agent systems (MASs), as shown in Figure 4. The figure depicts various factors that contribute to defining adaptive MASs, highlighting the interconnected nature of these systems and the role of cognitive capabilities, such as the Theory of Mind, in their functioning [5].

2.2.6 Theory of Emergent Communication

Emergent communication theory explains how communicative protocols may emerge naturally between interacting agents that have not been hardcoded. In AMACAI systems, this involves agents creating common symbols, codes, or conventions of language owing to recurring interactions. The emergent development of communication tools has facilitated flexible and effective coordination among agents, especially when conversations are dynamic or open-ended [12]. System robustness is also attributed to emergent communication because agents can self-organize their communication behavior when faced with changing goals or contexts.

2.2.7 Lifelong and Meta-Learning Theories

The idea of lifelong learning describes the capability of an agent to constantly learn through new experiences and use them without losing its previous knowledge [13]. Such an ability is essential in AMACAI systems that will be applied in realistic environments where user behavior and domain knowledge are expected to change over time. Meta-learning can help agents adapt to new tasks using little training data by exploiting knowledge acquired in past learning episodes. These theories enable conversational agents to grow in intelligence and personalization over time and enable them to cope with various conversations with minimal reprogramming.

2.3 Review Scope and Search Strategy

The review process was conducted on academic and technical publications published between 2015 and 2025 in fields associated with conversational AI, multi-agent systems, adaptive learning, and self-evolving architectures [14]. The notable fields of search are as follows:

- Architectural structures
- Centralized, decentralized, and modular systems in multiagent dialogue systems
- The use of techniques such as Hierarchical Reinforcement Learning (HRL), Partially Observable Markov Decision Processes (POMDPs), and memory-augmented models to address long-term conversations is also being explored.
- Adaptive methods include meta-learning, continual learning and emergent communication.

The domains of application include education, healthcare, customer service, virtual assistance, and cooperative AI.

Solutions at the hardware level, pure theoretical (no empirical data) models, and single-agent domain-specific systems are beyond the scope of this study.

2.4 Future Outlook and Open Challenges

Adaptive Multi-Agent Conversational AI is a research topic whose future is bright and is leaving the olden times of unchanging systems (single agents) to the times of changing multi-agent systems with advanced learning capabilities. Inspired by the theories of distributed cognition and lifelong learning, contemporary systems have the opportunity to be applied in the real world with superior memory and planning capabilities.

However, challenges remain in areas such as standard benchmarks, real-world implementation, comprehensive evalua- tion measures and system integration. This review provides a backdrop for future analyses of how architectural design, planning, and self-evolution properties affect the performance and adaptability of AMACAIs. The Methodology used in this review is described below.

3 Methodology

The proposed review follows a Systematic Literature Review (SLR) approach to provide a structured, transparent, and replicable method of identifying, appraising, and synthesizing the applicable academic work in the area of Adaptive Multi-Agent Conversational AI (AMACAI). The methodology addresses the overlap of three fundamental dimensions: architectural frameworks, long-horizon planning mechanisms, and self-evolving capabilities in conversational systems [15].

3.1 Search Strategy and Data Sources

A literature search was conducted in five significant academic databases covering computer science and artificial intelligence:

- IEEE Xplore
- ACM Digital Library (ACM DL)
- arXiv (preprints)
- ScienceDirect
- Google Scholar

The literature search was conducted on works published between 2015 and 2025, and new trends and recent developments in the sphere were sought. The keywords and their Boolean combinations were as follows:

- Multi-agent conversational AI
- Adaptive dialogue systems
- Dialog systems reinforcement learning AI self-evolution
- Conversational AI hierarchical planning

The titles, abstracts, and keywords were read to refine the search results and to obtain relevant information.

3.2 Scope and Limitations

Scope: This review concerns peer-reviewed academic literature, open-source frameworks, and benchmark assessments released between 2015 and 2025. This addresses three fundamental aspects: architectural design, long-horizon planning, and self-evolution of adaptive multi-agent conversational systems.

Limitations: It covers an extensive variety of use cases and approaches; however, the review does not provide a detailed low-level implementation and deployment at the domain-specific level or consider the scenario of dialogue systems.

3.3 Inclusion/Exclusion Criteria

The following inclusion criteria were applied to maintain the quality and relevance of the selected studies.

- Peer-reviewed journals or conference proceedings
- English language publications
- Studies involving adaptive Conversational AI, multi-agent systems, or self-evolving Conversational AI
- Articles that provide experimental or implementation evidence of a system

The exclusion criteria were as follows.

- Purely theoretical models not tested against data
- Rule-based dialog systems that lack adaptive or learning elements
- Redundant publications and popular literature

3.4 Categorization and Thematic Analysis

A thematic analysis approach was used to categorize and synthesize the findings [16]. All selected studies were evaluated and categorized into three fundamental dimensions.

- **Architecture:** Includes modular, centralized, and decentralized designs for multi-agent dialogue systems.
- **Planning:** The use of mechanisms such as Hierarchical Reinforcement Learning (HRL), Partially Observable Markov Decision Processes (POMDPs), and memory-augmented networks to deal with long-horizon dialogue has been investigated [17].
- **Self-Evolution:** Emphasis is placed on systems that use meta-learning, lifelong learning, emergent communi- cation, and autonomous adaptation to achieve this goal.

The papers were coded and charted into these categories to highlight the contributions, limitations, and future directions of the field.

Criteria Type	Inclusion	Exclusion
Timeframe	2015 – 2025	Publications before 2015
Language	English	Non-English papers
Topic Focus	Adaptive, multi-agent, planning, self-evolution in dialogue	Single-agent or rule-based dialogue systems
Source Type	Peer-reviewed journals, conferences, and preprints	Opinion articles, blogs, non-peer-reviewed reports
Empirical Evidence	Papers with experimental validation	Theoretical models without

		implementation or evaluation

Table 1: Inclusion and Exclusion Criteria

3.5 Data Extraction and Assessment Criteria

The data retrieved and reviewed in the literature were both quantitative and qualitative and included the following:

- Number of citations (to measure impact)
- Performance measures such as BLEU, ROUGE, perplexity, task success rates, and human evaluation scores
- Assessment procedures used in various studies
- Applicability and implementation status in the real world

Such systematic extraction permits comparisons among a wide variety of systems and methods [13].

Database	Keywords Used	Resulting Articles
IEEE Xplore	*"multi-agent conversational AI", "reinforcement learning in dia- logue"*	48
ACM Digital Library	*"adaptive dialogue systems", "emergent communication"*	56
arXiv	*"meta-learning for conversational agents", "continual learning"*	72
ScienceDirect	*"long-horizon dialogue planning", "hierarchical dialogue policies"*	34
Google Scholar	Combined queries from all above	100+

Table 2: Keyword Search and Database Mapping

3.6 Limitations

Although the SLR methodology pursues objectivity and comprehensiveness, some limitations have been identified in the literature.

- Bias in selection due to subjective interpretation during the screening stage
- The field is evolving quickly; therefore, the latest preprints or unpublished discoveries may be left out
- A restricted view of real-world performance may result from limited access to proprietary and industrial implementations

Despite these limitations, the methodology provides a solid foundation for analyzing and understanding trends and challenges in AMACAI and supports future research and system development [18].

4 Results and Analysis

This section synthesizes the findings reviewed in the literature and is organized into the following categories: archi- tectural evolution, planning capabilities, and self-evolution mechanisms in Adaptive Multi-Agent Conversational AI (AMACAI). It also includes a comparative overview of the performance indicators on benchmark datasets.

4.1 Architectural Trends

Architectural Trends

The architectural development of conversational AI over the past decade has evolved from static, rule-based systems to dynamic multiagent architectures underpinned by learning. Early dialogue managers, which relied on hardcoded templates or finite-state machines, lacked flexibility, scalability, and contextual integrity during extended interactions. In contrast, neural modular architectures—beginning with sequence-to-sequence and transformer-based systems—enable the division of labor among specialized agents for intent recognition, knowledge retrieval, response generation, and user engagement, resulting in situationally aware and adaptable dialogue.

The integration of Large Language Models (LLMs) into multi-agent dialogue systems marks a pivotal advancement. Researchers have begun adopting decentralized architectures, which improve fault tolerance and scalability, especially when LLM-based agents collaborate through shared memory structures or attention-based coordination protocols. For instance, AgentNet introduced a retrieval-augmented generation (RAG) framework for decentralized collaboration without central orchestration, allowing agents to dynamically specialize and route tasks in a DAG structure, thereby improving fault resilience and emergent collective behavior [19].

Likewise, transformer-based multi-agent models that share recurrent memory, such as the Shared Recurrent Memory Transformer (SRMT), pool individual agent memories into a global workspace, significantly enhancing coordination in tasks like multi-agent pathfinding and maze navigation compared to

Generation	Architecture Type	Key Features	Examples
Early (pre-2015)	Rule-based	Deterministic, static responses	ELIZA, ALICE
Intermediate (2015–2020)	Seq2Seq, Modular	Neural response generation, modular roles	Rasa, DialoGPT
Recent (2020–2025)	Multi-agent Transformer	Shared memory, decentralized coordination	ChatDev, CAMEL, AutoGen

Table 3: Architectural Evolution in Conversational AI

4.2 Planning Capabilities

Long-horizon planning remains a hallmark of sophisticated conversational agents (CAs). Flat policy models struggle with intermediary sub-goals and are prone to failure in long dialogues. In contrast, hierarchical planning architectures, such as Hierarchical Reinforcement Learning (HRL), enhance coordination in structured, multi-turn conversations by decomposing complex tasks into smaller dialogue segments. A recent experiment demonstrated that large language model (LLM) agents can spontaneously develop coherent communication norms through interaction, underscoring the effectiveness of high-level coordination in long-term dialogues [20].

High-level policies in hierarchical models specify global dialogue objectives or phases, whereas low-level policies manage immediate responses. This architecture maintains coherence over extended interactions such as customer onboarding, tutoring, and technical troubleshooting. Supporting this, the Hierarchical Neuro-Symbolic Decision Transformer couples a symbolic planner (for interpretable, globally consistent sub-goal sequencing) with transformer- based low-level policies, achieving significantly higher success rates and efficiency in long-horizon tasks compared to purely end-to-end neural models [21].

Some reviewed studies combine symbolic planners with neural policy agents to form hybrid systems that leverage structured reasoning and adapt to learning. These systems are especially valuable in negotiations, instructional conversations, and simulation-based training, where the alignment of goals and strategies is critical. The symbolic layer ensures interpretability and logical coherence, whereas the neural executor adds flexibility and adaptability, which are essential for managing unpredictable conversational environments [21].

Memory-augmented networks have also been applied to planning, where an agent can remember previous interactions and contextual changes to aid continuity and personalization across long periods. In multi-agent systems, planning is shared among agents; some agents may plan over strategic objectives, while others may plan in response to reactions or adapt their content to different users.

Recently, with the emergence of collaborative planning procedures in which agents exchange predictive models or planning results, turn-taking has become more fluent, redundant queries have been reduced, and task achievement has

Model Type	Planning Approach	Use Case	Performance Gain
Flat Policy Model	End-to-end RL	FAQ chatbots	Low
HRL-Based Multi-Agent Model	Hierarchical Reinforcement Learning	Tutoring systems, negotiation bots	High (↑30% suc- cess)
Neural-Symbolic Hybrid	Symbolic Planner + Neural Policy	Legal/medical ad- vising	Medium-High (↑20%)

Table 4: Planning Strategies Across Models

become more efficient. This has been particularly notable in active areas such as collaborative tutoring, healthcare advising, and virtual assistant systems, where the complexity of dialogue necessitates long-term and coordinated agent planning.

4.3 Evolution Mechanisms

Self-evolution, the ability to learn through experience, user feedback, and environmental changes, and adapt behavior to perform better in the future, is one of the hallmark goals of AMACAI systems. Three broad classes of mechanisms are cited in the literature as allowing this evolution: emergent communication, meta-learning, and self-correction through reinforcement learning.

Emergent communication is a mechanism used as the basis for multi-agent systems when agents design their protocols or languages for coordination purposes. Such emergent strategies prove particularly helpful where there are no pre-existing structures of communication or where such structures are inadequate (https://journals.sagepub.com/doi/10.3233/AIC-220147). Although emergent communication has been successful in enhancing collaboration and minimizing redun- dancy, it lacks interpretability; thus, it is difficult to diagnose agent behavior and debug faults.

Meta-learning algorithms allow agents to transfer knowledge about tasks done and quickly adapt to a new goal or a new environment with only a small amount of extra training. They have been observed to converge more quickly and generalize better than previous models, especially when the task dynamics (dialogue structures or user intentions) change significantly. Meta-learners: Adaptive task-switching and user profiling systems often use meta-learners to improve their performance.

Agents exhibiting self-correction properties can improve their behavior over time using feedback and reward signals, particularly in reinforcement learning environments. These models are characterized by slow but steady improvements in performance in areas such as task completion, user engagement and language fluency. Reinforcement signals. In multi-agent systems, reinforcement signals are occasionally shared or averaged among agents, encouraging team-level learning and alleviating competition.

Lifelong learning methods (that avoid catastrophic forgetting and integrate new knowledge) are also becoming popular. These models are used to ensure that innovative systems maintain their competencies and transform to address emerging challenges. When performed well, self-evolving systems exhibit increased robustness, situational awareness, and personalization in multiagent dialogue systems.

4.4 Quantitative Performance Summary

The results of a comparative analysis of performance on benchmark datasets, such as MultiWOZ, ConvAI2, ALFRED, and CRAFT, show that adaptive multi-agent systems have improved significantly compared with single-agent baselines. The success rates of tasks increased on average by 20-40 percent, and the coherence of dialogues and goal achievement were enhanced in multi-turn conversations. These systems are frequently judged by humans to be more relevant, personalized, and natural than traditional systems. Memory-augmented and planning-capable agents can continuously outperform flat models, particularly in long-horizon tasks.

4.5 Literature Gap

Although conversational AI has achieved remarkable advances, several underlying gaps exist that hinder the potential of adaptive multiagent systems [22]. These limitations intersect with the development of theories, applications, and

Dataset	Task Success Rate (Baseline)	Task Success Rate (AMACAI)	Human Satisfaction Rating
MultiWOZ	58%	80%	4.2 / 5
ConvAI2	62%	84%	4.5 / 5
CRAFT	49%	71%	4.0 / 5
ALFRED	55%	75%	4.3 / 5

Table 5: Benchmark Evaluation Summary

performance assessments. Such gaps and their subsequent work will lead to the future creation of robust and intelligent dialog systems that can operate in the dynamism of real-world reality [23].

4.5.1 Minimal set of subfields

Adaptive learning, multiagent coordination, and self-evolving architectures are typically performed in isolation. Most related efforts recognize and solve one or two of these aspects but never integrate architectural design, long-horizon planning, and self-evolution capabilities in an integrated manner [10]. Therefore, current systems lack the synergy required to simulate truly autonomous and context-sensitive conversational agents that can adapt to time and tasks effectively.

4.5.2 Evaluation measures

The quality of multiagent conversational systems is usually tested using conventional NLP scores, such as BLEU, ROUGE, and perplexity [12]. These measures are not indicative of the richness of multi-agent interactions, including factors such as agent communication, adaptability to user goals, and success in long-term planning. This deficiency de- mands multidimensional evaluation schemes that would allow quantifying the quality of the conversation, collaboration among agents, goal congruency, and real-time learning efficiency [24].

4.5.3 Deployment Problems in the Real World

The road to real-world deployment outside controlled settings is paved with important issues that afflict systems [25]. Problems encountered by these systems include computational scalability, lack of explainability, and lack of resilience to unforeseeable user interactions or beneficial ethics [26]. Furthermore, the lack of domain-specific adaptation limits their utilization in sensitive domains, including healthcare, legal services, and customer support.

4.5.4 Lack of Common Benchmarks

There is no widely accepted benchmark for evaluating AMACAI systems, which limits reproducibility and comparison [23]. This gap restricts reproduction, comparison, and general advancements in the field.

4.5.5 Underdeveloped Models of Self-Evolution

Although emergent communication and meta-learning show promise, their integration into real-time multi-agent dialogues remains underexplored [25]. Few systems demonstrate the dynamic ability to grow in ways that make them responsive to the continuing interaction of users, especially in long dialogues in which goals and contexts vary.

4.6 Future Implications in Domain-Specific Deployments

As adaptive multi-agent conversational systems advance, their use should spread into real-world high-stakes situations, including healthcare, law, education, and the arena of government, where the combination of planning, evolution, and multi-agent modularity is likely to be able not only to surmount complexities in the real world but also outsmart traditional single-agent systems [27] .

4.6.1 Healthcare

AMACAI systems can transform personalized care and clinical decision support in healthcare. Multi-agent dialogue systems with hierarchical planning and self-evolution can help in the triaging of symptoms, analysis of laboratory reports, and guiding patients through post-operative recovery. For example, one agent can be charged with analyzing symptoms, whereas the other can liaise with medical databases or follow-up. These systems would help alleviate delays in diagnosing the condition, particularly in rural or poor facilities [28].

Nonetheless, regulatory systems, including the Health Insurance Portability and Accountability Act (HIPAA) in the US and the General Data Protection Regulation (GDPR) in the EU, require high-quality data privacy, auditability, and explicability to protect patient information. To be compliant, multi-agent AI systems must either use Federated Learning or Differential Privacy methods, especially in situations where sensitive medical records must be dealt with on several nodes. Moreover, the upcoming EU Medical Device Regulation (MDR) and Software as Medical Device (SaMD) guidance by the FDA will be helpful in certifying AI agents as safe health-related devices [29].

4.6.2 Legal and Judicial Systems

Conversational AI systems can assist in the field of law with document summarization, statute interpretation, and legal aid navigation. Learning multi-agent systems that combine symbolic reasoning and learning (e.g., neural-symbolic hybrids) have the potential to assist with many tasks, such as checking for conflicts of interest, retrieving precedents, and auditing compliance. Such systems should be used in accordance with ethical AI principles—fairness, transparency, and accountability—as prescribed by frameworks such as the OECD AI Principles [30] and UNESCO's Ethics of Artificial Intelligence Recommendation [31].

Additionally, interpretability is prioritized by courts. Transparency and explainability are essential to ensure that stakeholders understand and can challenge AI-generated outcomes, particularly in high-stakes contexts such as legal proceedings [32, 33]. Furthermore, HCI design concepts such as eXplainable AI (XAI)—which focuses on making AI reasoning intelligible—and accountable conversational UX should be integrated, allowing legal professionals to verify results with confidence [34, 35].

4.6.3 Education and Tutoring

Adaptive systems have been used in education to personalize learning paths, with various agents responsible for curriculum planning, engagement and feedback. Subordinate learning agents can adapt to students' performance by dynamically arranging long-term educational goals. Effective adaptive tutoring agents must adhere to human-centered design principles to ensure usability, accessibility, and support for diverse learners, as outlined in [36]. Furthermore, the IEEE 7000 series—especially the IEEE 7000-2021 standard on ethical system design—provides a structured framework for embedding autonomy, equity, and ethical considerations into AI systems for children and disadvantaged students [37].

4.6.4 Ethical and Societal Considerations

Ethical issues surrounding AMACAI systems are becoming increasingly prominent, amplified by challenges such as distributed responsibility, unpredictable behavior, and moral coordination inherent in multi-agent architectures. Ensuring system-level explainability is critical, not only for effective debugging but also to build trust among human stakeholders [38, 39].

Recent HCI recommendations, including those from the AI Now Institute and the ACM FAccT community, emphasize human-in-the-loop control, ontological transparency, and the importance of post-deployment surveillance in multi-agent systems [40, 41].

Without periodic ethical testing, particularly in simulated conditions, emergent communication or reinforcement learning-based agents risk deviating from acceptable norms through reward-hacking or undesired behaviors during self-organization [38].

5 Discussion

The findings of this review support the concept of the blistering rate of development and the growing sophistication of Adaptive Multi-Agent Conversational AI (AMACAI) systems. Through the abuse of architectural modularity, long-horizon planning, and self-evolving mechanisms, these systems are transforming the operation of conversational agents in various domains. However, these developments have had a fair share of serious trade-offs, challenges, and other implications that are equally important to discuss critically.

Architecture and planning are among the synergies that have emerged in this domain. Modular multi-agent systems allow the assignment of planning tasks to specially designated modules. This enables agents to become specialists in single tasks, such as intent recognition and strategic goal generation, without imposing a burden on the central controller of the system. Such a separation of concerns establishes superior scalability, flexibility, and performance in dynamic conversational environments. However, this modularity may also cause coordination problems, that is, in instances when agents possess interdependent goals and do not enjoy an integrated representation of the global dialogue context.

Adaptive agents identify themselves through interactions because the interactions are personalized on a per-request basis. However, this flexibility in the behavior of an agent cannot be accepted in mission-critical applications, such as the healthcare domain or the defense industry. Predictability and control are paramount in these fields. In addition, reinforcement/meta-learning may optimize poorly defined agents for undesirable behaviors, particularly in open-ended environments. This raises concerns regarding misalignment, where the agent maximizes incorrect goals, thereby jeopardizing safety and integrity of the system.

Centralized and decentralized coordination also have trade-offs. Centralized systems tend to exhibit better global alignment and coherence, particularly in task-based dialogue. However, they suffer from bottlenecks and scalability issues. In contrast, decentralized systems enable autonomy and parallelism among agents but can provoke the destruction of dialogue streams and non-consistent user experiences if synchronization primitives are weak or fail.

Adaptive multi-agent systems are highly data intensive. Their performance is typically defined by access to large and diverse datasets or high-resolution simulations. This makes them expensive to train and limits their application in data-scarce environments such as space. In addition, despite the progress made in the fields of meta-and transfer learning, domain generalization remains limited. Most systems are strongly optimized for a specific set of environments and require extensive retraining or domain adaptation to function in different domains.

The evaluation of these models remains challenging. Standard NLP metrics do not reflect the complexity of multiagent dialogues, particularly in terms of the collaboration quality, adaptivity, and long-horizon coherence. Human evaluation is the best standard to date; however, it is resource intensive and non-reproducible. In addition, there is a lack of benchmarks that explicitly focus on assessing systems with long horizon planning and self-evolution.

The question of ethics and safety casts a large shadow over the systems that can be developed. Emergent behaviors can cause goal drift or manipulate the reward signals. The danger with such systems is that they can easily pass the boundary of user trust or ethics without restrictions. As technologies of this sort approach the stage of practical application, there is a need to ensure transparent decision-making, accountability, and safe learning processes in their use.

AMACAI systems have extensive implications for various sectors. They can offer individual tutoring in education, patient monitoring, and interaction in health care. They assert that scale, awareness, and communication are required for effective customer care and protection. However, to keep this promise, future innovations must be accompanied by moral responsibility, interpretability, and good-quality criteria.

6 Conclusion and Recommendation

Conclusion

This review systematically surveys the current state-of-the-art, foundational components, and prospective research directions in Adaptive Multi-Agent Conversational AI (AMACAI). The analysis foregrounds the architectural evolution of these systems, their long-horizon planning capabilities, and the emergence of self-evolving properties within next-generation conversational agents.

AMACAI research signifies a paradigm shift from inflexible, rule-based architectures to contextually aware, dynamically adaptable agents capable of collaborative and autonomous behavior. Recent advancements have transformed monolithic conversational frameworks into modular and distributed multiagent architectures, notably leveraging transformer-based specialization, coordination, and scalability. These developments have markedly enhanced the robustness and task orientation of conversational agents, enabling them to operate effectively in complex, real-world environments.

Hierarchical reinforcement learning (HRL) and memory-augmented networks have emerged as critical enablers for sustained long-term goal management and continuity in multi-turn dialogues. The integration of symbolic reasoning further enhances logical consistency, whereas meta-learning and continual learning frameworks equip agents with the ability to generalize from limited data and adapt continuously to novel scenarios. Meanwhile, emergent communication has allowed agents to develop new behavioral strategies beyond explicit programming.

Despite these advances, several challenges persist. Resource allocation remains a significant obstacle to achieving resilient and scalable AMACAI deployment. Current evaluation paradigms inadequately capture the adaptability and collective efficacy of multiagent systems during extended interactions. Additionally, safety and value alignment are becoming increasingly pressing issues as agents exhibit behaviors that deviate from their original design intent, sometimes resulting in unanticipated or ethically ambiguous emergent phenomena.

Another impediment is the absence of standardized protocols for inter-agent communication and data structuring, which restricts interoperability, reproducibility, and comparative benchmarking across AMACAI implementations. These limitations collectively highlight the necessity for further foundational work to enable the development of robust, explainable, and ethically aligned AMACAI systems in the future.

Recommendations

To address the identified challenges and catalyze progress in AMACAI, the following recommendations are proposed.

- **Establish Domain-Specific Evaluation Criteria:** Develop comprehensive, context-sensitive metrics and benchmarks tailored to the long-horizon and adaptive nature of multi-agent conversational systems.
- **Foster Interdisciplinary Collaboration:** Promote joint research efforts spanning artificial intelligence, cognitive science, ethics, and human-computer interaction to ensure that AMACAI systems are socially compatible and ethically grounded.
- **Advance Data-Efficient Learning Paradigms:** Leverage transfer learning and few-shot learning strategies to reduce the data and computational requirements of training scalable, adaptable agents.
- **Enhance Transparency and Explainability:** Design communication structures and decision-making frame- works that are inherently interpretable, thereby fostering user trust and system accountability.
- **Align with Human Values:** Integrate symbolic reasoning and neural plasticity-inspired mechanisms to ensure that agent behaviors align with societal norms and user expectations.
- The implementation and sustained advancement of these recommendations are critical for realizing scalable, trustworthy, and human-centric AMACAI systems in the future.

References

1 J. Li, M. Zhang, N. Li, D. Weyns, Z. Jin, and K. Tei. Generative ai for self-adaptive systems: State of the art and research roadmap. *ACM Transactions on Autonomous and Adaptive Systems*, 19(3):1–60, 2024.

2 Z. Tao, T.E. Lin, X. Chen, H. Li, Y. Wu, Y. Li, Z. Jin, F. Huang, D. Tao, and J. Zhou. A survey of the self-evolution of large language models, 2024.

3 T. Guo, X. Chen, Y. Wang, R. Chang, S. Pei, N.V. Chawla, O. Wiest, and X. Zhang. Large language model-based multi-agents: A survey of progress and challenges, 2024.

4 Z. Chu, Y. Wang, F. Zhu, L. Yu, L. Li, and J. Gu. Professional agents: Evolving large language models into autonomous experts with human-level competencies, 2024.

5 N. Nezamoddini and A. Gholami. A survey of adaptive multi-agent networks and their applications in smart cities.

6 *Smart Cities*, 5(1):318–347, 2022.

7 C. Zhang, S. He, J. Qian, B. Li, L. Li, S. Qin, Y. Kang, M. Ma, G. Liu, Q. Lin, and S. Rajmohan. Large language model-brained gui agents: A survey, 2024.

8 Omdia. Why rigorous definitions matter in the agentic ai conversation, 2025. Accessed: 2025-07-02.

9 B. Liu, X. Li, J. Zhang, J. Wang, T. He, S. Hong, H. Liu, S. Zhang, K. Song, K. Zhu, and Y. Cheng. Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems, 2025.

10 A. Kalyuzhnaya, V. Vorona, N. O. Nikitin, N. Chichkova, K. Fatkhiev, S. Mityagin, Y. Aksenkin, A. Boukhanovsky,

11 K. Fedorin, E. Lutsenko, and A. Getmanov. Llm agents for smart city management: Enhancing decision support through multi-agent ai systems, 2025.

12 J. Liao, M. Wen, J. Wang, and W. Zhang. Marft: Multi-agent reinforcement fine-tuning, 2025.

13 Z. Feng, R. Xue, L. Yuan, Y. Yu, N. Ding, M. Liu, B. Gao, J. Sun, and G. Wang. Multi-agent embodied ai: Advances and future directions, 2025.

14 G. Liu, P. Zhao, L. Liu, Y. Guo, H. Xiao, W. Lin, Y. Chai, Y. Han, S. Ren, H. Wang, and X. Liang. Llm-powered gui agents in phone automation: Surveying progress and prospects, 2025.

15 Y. Du, W. Huang, D. Zheng, Z. Wang, S. Montella, M. Lapata, K.F. Wong, and J.Z. Pan. Rethinking memory in ai: Taxonomy, operations, topics, and future directions, 2025.

16 S. Hu, T. Huang, G. Liu, R.R. Kompella, F. Ilhan, S.F. Tekin, Y. Xu, Z. Yahn, and L. Liu. A survey on large language model-based game agents, 2024.

17 Y. Liu, X. Cao, T. Chen, Y. Jiang, J. You, M. Wu, X. Wang, M. Feng, Y. Jin, and J. Chen. A survey of embodied ai in healthcare: Techniques, applications, and opportunities, 2025.

18 W. Liu, J. Qin, X. Huang, X. Zeng, Y. Xi, J. Lin, C. Wu, Y. Wang, L. Shang, R. Tang, and D. Lian. The real barrier to llm agent usability is agentic roi, 2025.

19 J. Yu, Y. Qin, H. Che, Q. Liu, X. Wang, P. Wan, D. Zhang, and X. Liu. Position: Interactive generative video as a next-generation game engine, 2025.

20 A. Chen, Y. Wu, J. Zhang, S. Yang, J.T. Huang, K. Wang, W. Wang, and S. Wang. A survey on the safety and security threats of computer-using agents: Jarvis or ultron?, 2025.

21 Yingxuan Yang, Huacan Chai, Shuai Shao, Yuanyi Song, Siyuan Qi, Renting Rui, and Weinan Zhang. Agentnet: Decentralized evolutionary coordination for llm-based multi-agent systems. arXiv preprint, 2025. Introduces a decentralized RAG-based DAG coordination framework.

22 Ariel Flint Ashery, Andrea Baronchelli, et al. Emergent social conventions and collective bias in llm populations.

23 *Science Advances*, 2025. Demonstrates spontaneous emergent communication among agent populations.

24 Ali Baheri and Cecilia O. Alm. Hierarchical neuro-symbolic decision transformer. arXiv preprint, 2025. Combines symbolic planning with transformer-based policies for long-horizon tasks.

25 S. Du, J. Zhao, J. Shi, Z. Xie, X. Jiang, Y. Bai, and L. He. A survey on the optimization of large language model-based agents, 2025.

26 J. Zheng, C. Shi, X. Cai, Q. Li, D. Zhang, C. Li, D. Yu, and Q. Ma. Lifelong learning of large language model-based agents: A roadmap, 2025.

27 Z. Wang, K. Wang, Q. Wang, P. Zhang, L. Li, Z. Yang, K. Yu, M.N. Nguyen, L. Liu, E. Gottlieb, and M. Lam. Ragen: Understanding self-evolution in llm agents via multi-turn reinforcement learning, 2025.

28 F. Tang, H. Xu, H. Zhang, S. Chen, X. Wu, Y. Shen, W. Zhang, G. Hou, Z. Tan, Y. Yan, and K. Song. A survey on

29 (m) llm-based gui agents, 2025.

30 Z.Z. Li, D. Zhang, M.L. Zhang, J. Zhang, Z. Liu, Y. Yao, H. Xu, J. Zheng, P.J. Wang, X. Chen, and Y. Zhang. From system 1 to system 2: A survey of reasoning in large language models, 2025.

31 Y. Deng, Y. Li, B. Ding, and W. Lam. Leveraging long short-term user preference in conversational recommenda- tion via multi-agent reinforcement learning, 2023.

32 L. Tudor Car, D. A. Dhinagaran, B. M. Kyaw, T. Kowatsch, S. Joty, Y.-L. Theng, and R. Atun. Conversational agents in health care: Scoping review and conceptual analysis, 2020.

33 O. Fayayola, O. Olorunfemi, and P. Shoetan. Data

privacy and security in it: A review of techniques and challenges, 2024.

34 OECD. Oecd ai principles: Promoting trustworthy ai. https://oecd.ai/en/ai-principles, 2024. Accessed July 2025.

35 UNESCO. Recommendation on the ethics of artificial intelligence. https://www.unesco.org/en/ artificial-intelligence/recommendation-ethics, 2021. Accessed July 2025.

36 OECD. Transparency and explainability (principle 1.3). https://oecd.ai/en/dashboards/ ai-principles/P7, 2024. Accessed July 2025.

37 UNESCO. Ai and the judiciary: Balancing innovation with integrity. https://www.unesco.org/en/ articles/ai-and-judiciary-balancing-innovation-integrity, 2025. Accessed July 2025.

38 Alejandro Barredo Arrieta et al. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. https://arxiv.org/abs/1910.10045, 2019. Accessed July 2025.

39 Michael Nasir and Abdul Bamako. Transparent justice: Legal frameworks for explainable ai in law. https://www.researchgate.net/publication/392312198_Transparent_Justice_Legal_ Frameworks_for_Explainable_AI_in_Law, 2025. Accessed July 2025.

40 ISO/TC 159/SC 4 Ergonomics of human–system interaction committee. Iso 9241-210:2019 – ergonomics of human–system interaction – part 210: Human-centred design for interactive systems. International Standard, 2019. Prepared by ISO/TC 159/SC 4; withdrawal of ISO 13407.

41 IEEE. Ieee 7000-2021 – model process for addressing ethical concerns during system design. IEEE Standard, 2021. Published by IEEE; authored by committee.

42 Alan Chan, Rebecca Salganik, Alva Markelius, Chris Pang, Nitarshan Rajkumar, Dmitrii Krasheninnikov, Lauro Langosco, Zhonghao He, Yawen Duan, Micah Carroll, Michelle Lin, Alex Mayhew, Katherine Collins, Maryam Molamohammadi, John Burden, Wanru Zhao, Shalaleh Rismani, Konstantinos Voudouris, Umang Bhatt, Adrian Weller, David Krueger, and Tegan Maharaj. Harms from increasingly agentic algorithmic systems. *arXiv*, 2023. Discusses distributed responsibility and emergent harms in agentic systems.

43 Stephanie Baker and Wei Xiang. Explainable ai is responsible ai: How explainability creates trustworthy and socially responsible artificial intelligence. *arXiv*, 2023. Connects XAI with system-level trust and explainability.

44 AI Now Institute. Ai now 2018 report. Report, 2018. Recommends human-in-loop control and full-supply-chain transparency in AI systems.

45 ACM FAccT Community. Acm facct 2025 focus areas. Conference Position Statement, 2025. Highlights importance of HCI, decision support, human-in-loop, and surveillance practices.